Vegetation Resources Inventory

Sample Selection Procedures for Ground Sampling

Prepared by
Ministry of Sustainable Resource Management
Terrestrial Information Branch
for the Terrestrial Ecosystem Task Force
Resources Information Standards Committee

December 2002

Version 3.3

© The Province of British Columbia
Published by the
Resources Information Standards Committee

National Library of Canada Cataloguing in Publication Data

Main entry under title:

Vegetation resources inventory. Sample selection procedures for ground sampling. – Version 3.3. –

Cover title.

"December 2002"

Also available on the Internet.

ISBN 0-7726-4891-3

1. Vegetation surveys - British Columbia – Methodology – Handbooks, manuals, etc. 2. Forest surveys - British Columbia - Methodology – Handbooks, manuals, etc.

3. Environmental sampling – Handbooks, manuals, etc.

I. British Columbia. Ministry of Sustainable Resource Management. Terrestrial Information Branch. II. British Columbia. Resources Information Standards Committee. III. Title: Sample selection procedures for ground sampling.

QK203.B7V434 2002 333.95'311'09711 C2002-960264-5

Additional Copies of this publication can be purchased from:

Government Publication Services
Phone: (250) 387-6409 or
Toll free: 1-800-663-6105
Fax: (250) 387-1120
www.publications.gov.bc.ca

Digital Copies are available on the Internet at:
http://srmwww.gov.bc.ca/risc

Abstract

This report outlines the components of the Vegetation Resources Inventory (VRI) sample selection process. These components include:

- sampling method,
- assembling the population,
- stratification,
- selecting sample polygons,
- selecting sampling points within polygons,
- quality assurance of the selection process,
- documentation, and
- approval protocols.

This document is intended for use by technical persons with an understanding of the VRI process and experience in data manipulation.

Acknowledgments

The Government of British Columbia provides funding of the Resources Information Standards Committee work, including the preparation of this document. The Resources Information Standards Committee supports the effective, timely and integrated use of land and resource information for planning and decision making by developing and delivering focussed, cost-effective, common provincial standards and procedures for information collection, management and analysis. Representatives to the Committee and its Task Forces are drawn from the ministries and agencies of the Canadian and the British Columbia governments, including academic, industry and First Nations involvement.

The Resources Information Standards Committee evolved from the Resources Inventory Committee which received funding from the Canada-British Columbia Partnership Agreement of Forest Resource Development (FRDA II), the Corporate Resource Inventory Initiative (CRII) and by Forest Renewal BC (FRBC), and addressed concerns of the 1991 Forest Resources Commission.

For further information about the Resources Information Standards Committee, please access the RISC website at: http://srmwww.gov.bc.ca/risc/.

Terrestrial Ecosystems Task Force

The Vegetation Inventory Working Group was formed in 1993 and issued their final report in March 1995 on a "Proposed New Inventory" for British Columbia. The Ministry of Forests, Resources Inventory Branch, in cooperation with the Ministry of Environment and other Ministry of Forests branches and consultants, developed the suite of Vegetation Resources Inventory Procedures based on the recommendations in that report. Many individuals were involved in writing the original version of the various Vegetation Resources Inventory Procedures documents.

For questions concerning the content of this publication please contact the:

Vegetation Resources Inventory Coordinator
c/o Ministry of Sustainable Resources Management
Terrestrial Information Branch
PO Box 9993, 722 Johnson Street
Victoria, British Columbia
V8W 9C2

Phone: (250) 387-1314
Fax: (250) 387-5999

Table of Contents

1.0 Introduction 1
1.1 Background 1
1.2 Target audience 1
1.3 Purpose of the document 1
2.0 Assembling the Population Data 3
2.1 Establishing data sources 3
2.2 Processing source files 3
2.3 Reconciling graphics and attribute data 3
2.4 Combining data from different sources 4
2.5 Assigning overlay attributes to forest cover polygons within the population boundaries 4
2.6 Administrative boundaries defining the population 4
2.7 Checking for obvious errors 5
2.8 Identifying the population for sampling 5
3.0 Selecting sample polygons 7
3.1 Principles 7
3.2 Process 7
3.2.1 Land cover categories 7
3.2.2 Vegetated treed category 8
3.3 Requests for additional samples 9
4.0 Selecting Sampling Points Within Polygons 11
4.1 Principles 11
4.2 Process 11
4.3 Small polygons 12
5.0 Quality Assurance 13
5.1 Principles 13
5.2 Checking the population list 13
5.3 Checking sample distributions 13
5.4 Evaluate data gaps 14
5.5 Check for effect of majority rule 14

6.0 Documentation 15
6.1 Purpose 15
6.2 Data versions 15
6.3 Features of selection process 15
6.4 Stratification criteria 15
6.5 Sample allocation 16
6.6 Sample size 16
6.7 Sample and strata weights 16
6.8 Documentation of errors 16
6.9 Use of sample lists 16
6.10 Archiving of data 16
7.0 Reporting 19
7.1 Purpose 19
7.2 Report contents 19
8.0 Approval Protocols 21
8.1 Overview 21
8.2 Validation of population 21
8.3 Validation of selected samples 21
8.4 Approval of final report 21
Appendix 1 - Glossary 23
Appendix 2 - Sample List Attributes 25

1.0 Introduction

1.1 Background

The Vegetation Resources Inventory (VRI) is designed to answer two questions:

1. Where is the resource located?
2. How much of a given vegetation resource (for example, timber or coarse woody debris) is within an inventory unit?

The VRI is carried out in two phases. The first phase, photo-interpretation, involves estimating vegetation polygon characteristics from existing information, aerial photography, or other sources. No formal statistically based sampling is done in the photo-interpretation phase. The second phase, statistically based ground sampling, provides the information necessary to determine how much of a given characteristic is within the inventory area.

The ground measurements resulting from the ground sampling process are used to estimate means and totals for the population. The relationship between the polygon estimates and ground samples is then used to adjust the photo-interpreted polygon estimates.

The goal of ground sampling is to obtain reliable data from a small percentage of the population to make certain inferences about that population. Errors in the sample selection process are translated directly into errors in the resultant data.

1.2 Target audience

This document provides procedures to assist professional foresters and data management technicians in selecting statistically valid samples either for the VRI or for other purposes where the sampling design discussed in this document is applicable. No advanced statistical knowledge is required to apply these procedures; however, users should familiarize themselves with some of the statistical issues discussed to avoid misapplication of the procedures.

Those who use this manual in the sample selection process are advised to seek the assistance of a qualified biometrician to review the sample selection in any VRI project. It is also recommended that a VRI Strategic Inventory Plan (VSIP) (http://srmwww.gov.bc.ca/tib/veginv/project_management.htm) precede any sampling process where the procedures described in this document are used. The VSIP identifies high level forest management issues. It does not provide the project level detail contained in the Vegetation Inventory Pre-Inventory Plan (VPIP).

1.3 Purpose of the document

The sample selection process is critical to the resulting accuracy of the sample data in making inferences regarding the target population.

This document is written to ensure that statistical rigor is maintained at all times in the VRI sample selection processes. More specifically, the document is designed to:

1. Ensure that all population elements eligible for selection are included in the sample selection database (or sampling frame).

2. Ensure that all elements have a known chance (probability) of selection.
3. Eliminate subjective choice of samples.
4. Ensure the selection process is repeatable and auditable.
5. Ensure sample selection conforms to VRI sampling objectives and design as defined in the VPIP.
6. Identify standard outputs (sample list, QA reports, and sample section reports).
7. Identify archiving procedures.

This document assumes that:

1. Either the photo-interpreted information was collected to VRI standards, including the BC Landcover Classification Scheme (BCLCS) data (see Vegetation Resources Inventory BC Landcover Classification Scheme Volume, March 30, 2001, version 1.2), or, alternatively, that the photo-interpretation data to be used for sample selection contains a minimum of required attributes. In the latter case, the BCLCS would be derived.
2. The attributes required to pre-stratify the land are reasonably accurate.
3. The VPIP has clearly identified the target population, sampling objectives, sample sizes, and criteria for pre-stratification and sample allocation.
4. The selection of sample polygons for ground sampling is based on stratified sampling using the Probability Proportional to Size With Replacement (PPSWR) sampling method.

2.0 Assembling the Population Data

2.1 Establishing data sources

The sample selection process starts with the identification of sources of attribute and graphics data for the target population as defined by the sampling objectives.

Some of the information may be available within Ministry of Sustainable Resources Management (MSRM) databases. Other information may be located in licensee offices or under private management. If so, these agencies must all be approached to provide the information.

If the data are not available, either a process for improvising the data should be stipulated or the area in question should be eliminated from consideration, thus reducing the size of the population being sampled.

2.2 Processing source files

Most of the data used for VRI sample selection has two components:

- **Attribute files** - containing numeric or descriptive data.
- **Graphics files** - digital information showing polygons delineating areas of "similar" attributes, such as forest cover maps and recreation maps.

Although the main interest is initially in forest cover information, other ancillary data such as biogeoclimatic (BEC) zones and operability lines may be necessary for stratification purposes.

It is often most efficient to not only acquire all information required to select samples, but also to include information that will aid in project management and analysis of sample data when preparing the preparation sample list. A detailed listing of those attributes that are both required and recommended for sample selection and project management are included in Appendix 2.

Care must be taken to note the versions of data used when compiling information to form the target population as each data source undergoes periodic upgrading and updating. This can result in a population that differs between that used for sample selection and that used for analysis, sometimes at a much later date.

2.3 Reconciling graphics and attribute data

The graphics and attribute data should be edited and assembled simultaneously once the gathering of source information is complete.

All delineated polygon areas in the graphics file should have an identification number and descriptive information indicating what type of land cover is represented by the polygon and a valid area value.

The polygons should be checked to ensure that the population is complete and that all valid polygons were included.

2.4 Combining data from different sources

Gaps in the data will occur where maps or administrative boundaries from various sources do not meet precisely. In these cases, there will be no representation for these areas in the sample selection database. Overlapping data will result in double counting of the same areas in the sample selection database.

Gaps do not have any associated descriptive information and are difficult to process during sample selection and data analysis. This is because the gap areas are real on the ground, but have no defining features on the map. Where data are not available, they must be accounted for by either:

- Providing temporary map and attribute characteristics to the gap areas to allow completion of the selection process, or
- Excluding the information from the population. This should only be done when the magnitude of the missed areas are known and considered a significant potential bias.

Overlapping data should be identified and the source of information considered to be most reliable should be used in the assembly of the sample selection database.

Decisions made on how missing and overlapping data are accounted for in the assembly of the sample selection database should be well documented.

2.5 Assigning overlay attributes to forest cover polygons within the population boundaries

Population maps and their associated attribute data are formed by merging different graphic overlays that contain characteristics such as BEC zones, operability lines and forest cover polygons. As a result, the forest cover polygons are split into smaller segments called resultant polygons.

Data processing and long-term handling is complicated by the presence of the numerous resultant polygons being associated with a single polygon. To eliminate this resultant problem, the following majority rules have been developed:

1. Evaluate the number of resultants generated from a candidate polygon.
2. Determine the area sum of the resultant polygons by overlay attribute for each polygon.
3. Assign the overlay attribute that makes up the greatest percentage area of the polygon to that polygon.

In some cases, such as operability, a value of 'operable' may be assigned to all polygons that have any portion of the polygon overlaid by an operable value. This decision would ensure that all potentially operable polygons are included in a sample selection that emphasizes operability. However, this process may be flexible provided the rules are understood and well documented.

2.6 Administrative boundaries defining the population

If a target population is defined to include all sample elements within an inventory unit or sub-unit, it may be necessary to define the administrative boundary of the target population explicitly in order to split any polygons along this boundary into resultant polygons. Such administrative boundaries include district, TSA, TFL or private land boundaries. The explicit definition of the population boundary makes it possible to ensure that:

1. All polygons belonging to target population are present in the population frame.

2. All areas outside the target area are excluded from the list.
3. Gaps are described and accounted for.
4. The population area based on summing individual polygon areas matches against the known Gazetted area for the inventory unit or sub-unit.

2.7 Checking for obvious errors

During data assembly it is necessary to undertake some level of quality control to ensure that the resulting sampling frame accurately reflects the target population.

The sheer magnitude of the population may hinder the thorough checking of the list for potential errors. In addition, it is often difficult to determine which of the key attributes in stratifying the list may have data entry or data management error. As a consequence, error checking is usually limited to the following operations:

1. Look at general area distributions, such as total polygon area, polygon area by biogeoclimatic zone, polygon area by leading species, or polygon area by operability for the population and compare it to expected results.
2. Where possible, plot colour-themed maps showing the distribution of forest cover attributes and compare these distributions against other overlay values, such as biogeoclimatic zones. For instance, one might expect a general species, age, height, crown closure, and history distribution to coincide with certain biogeoclimatic zones or administrative units.

A more thorough review of the population and how well the sample drawn relates to the population is described in Section 5.

2.8 Identifying the population for sampling

When satisfied that all sources of data have been assembled correctly, and that any obvious data source errors have been identified and accounted for, the identification of the population for sampling may require that some portion of the data be netted out. For instance, the population may be limited to the vegetated treed, operable portion of the TSA, excluding all private, federal, and park lands.

It is very important in these cases to identify the forest cover attributes used to identify the vegetated treed portion of the landbase, the operability values included, and the administrative (ownership) values included in the final population. In cases where the populations are defined in a very complex manner, such as administrative units used in timber supply analyses, documentation of how the populations are derived is critical for future analysis and adjustments.

3.0 Selecting sample polygons

3.1 Principles

Once polygons in the population have been stratified based upon the criteria specified in the VRI Project Inventory Plan, the selection of sample polygons is implemented within each stratum in two stages:

1. Polygon selection is done using the probability proportional to size with replacement (PPSWR) sampling method. Selecting polygons with PPSWR means that the larger polygons have a proportionately greater chance of being included in the sample relative to smaller polygons. As the PPSWR sampling design requires the replacement of selected polygons back into the sampling frame, sample polygons may be selected more than once.
2. Select a point or points within a sample polygon at random and with replacement (see Section 4).

With this methodology, each sample selection is totally independent of the other selections.

3.2 Process

3.2.1 Land cover categories

Pre-stratification is a requirement of the probability proportional to size with replacement (PPSWR[1]) sampling strategy. With pre-stratification, PPSWR is more efficient than simple random sampling. For most inventory units in BC, it is recommended that the photo-interpreted polygons be stratified into the following domains or categories:

1. non-vegetated,
2. vegetated non-treed, and
3. vegetated treed polygons.

The non-vegetated polygons include lakes, rock, bare ground and double line roads which have no vegetation cover. The application of PPSWR in this category does not require further sub-stratification.

The vegetated non-treed category consists of polygons with vegetation but with no trees. Where trees do exist, they constitute less than 10% of the vegetation crown closure. As in the case for the non-vegetated category, further sub-stratification will likely not be required.

In situations where full VRI sample selection is required, where all the three land cover categories should be sampled, the full VRI sample may be allocated proportionally to the three domains.

---

[1] For further information on PPSWR, please refer to 'The Statistical Estimation and Adjustment Process Using PPSWR Sampling Design in the Vegetation Resources Inventory' (http://srmwww.gov.bc.ca/tib/veginv/techdocs/techdocuments.htm)

3.2.2 Vegetated treed category

The vegetated treed category consists of polygons with forest cover on them. By definition treed polygons have 10% tree crown closure or greater. For timber emphasis plots (TEPs), all the samples will be within the vegetated treed category. The application of PPSWR in this category requires further sub-stratification, and should be carried out as follows:

1. **Separate the polygons by strata.**

   Review the stratification criteria identified in the VPIP to ensure that polygons are classified into appropriate categories. Strata will typically be defined by leading species, or species groups, and then further stratified by a three or four volume classes, or appropriate surrogates. For example, the landbase may be stratified into 4 leading species groups, or based on a combination of leading species and growth or economic characteristics.

   It is recommended that the number of strata be restricted to range between 2 and 6, as the benefits of stratification are lost if too many strata are created. Also, if a minimum sample size is required for the smallest stratum, sample size must increase substantially as the number of strata increase if proportional allocation is to be maintained.

2. **Separate the strata into sub-strata.**

   For efficiency, and to ensure an even distribution of samples across stratum, we recommend creating 3 sub-strata per stratum based on equal numbers of polygons. The attribute on which the sub-stratification should be based is volume, however, in the absence of volume, a surrogate attribute such as basal area can be used.

3. **Determine the sample allocation (number of sampling points) for each stratum.**

   This information is usually partially specified in the VPIP. Currently, a minimum sample size of 15 samples per stratum is recommended, but not necessarily required. The overall sample size and the allocation to the strata are predetermined prior to sample selection and is usually specified in the VSIP.

4. **Determine the total area size for each sub-strata within a stratum.**

   The sub-strata should be designed to result in sub-groupings (within a stratum) with an approximately equal number of polygons. Allocation of samples to sub-strata should be proportional to total area of each sub-stratum.

5. **Accumulate the individual polygon areas beginning with the first polygon in the sub-stratum to the last polygon in each sub-stratum.**

   These accumulated areas should be stored as a new attribute, and will be used in selecting the sample.

6. **Generate as many random numbers as there are allocated samples for each sub-stratum.**

   The random numbers should be generated separately for each sub-stratum, and should range in size between 0 (zero) and the total area of each sub-stratum.

7. **Use the generated random numbers to identify the sample polygons.**

   A polygon is selected if a generated random number is larger than the accumulated total area corresponding to the polygon immediately preceding it, and the random number is smaller than or equal to its accumulated area.

8. **A polygon may be selected more than once.**

   After the first selection, the polygon is retained in the list and is eligible for any subsequent selection. The number of sample clusters that are established in a sample polygon should be equal to the number of times the polygon was selected during the selection process.

   Combine all samples from all strata into a single list when completed. Randomize their order, assign unique sample numbers (that must not change). Currently MSRM assigns a project id to these samples. A copy of the sample list should also be appended to either the VSIP or the VPIP.

   It would also be advisable at this point to select additional sample polygons (e.g., 50% per stratum) for future use.

   The mechanics of the sample selection process will be unique to each project, therefore, it would be advisable to consult with a biometrician.

3.3 Requests for additional samples

Additional samples may be needed to increase sample size. These can be selected using the PPSWR approach outlined in Section 5.2 at any time during field data collection, provided the population is not changed. Complete documentation is critical to ensure project integrity.

Suggested documentation:

1. Name of the person placing the request.
2. Date the request is made.
3. Project ID.
4. Number of samples requested.
5. Reasons why the additional sample are required.
6. Time frame for supplying the sample list.

This documentation is required should it becomes necessary to track the history of the sample selection process.

4.0 Selecting Sampling Points Within Polygons

4.1 Principles

The objective is to establish a random point within the polygon selected. Determination of a sample location using a 100 meter grid (described below) is often practical whether using digital or hard copy maps (with grids plotted). GIS techniques to determine random points within a polygon are also acceptable.

A 100-meter provincial grid has been established for selecting sampling points within polygons. This grid was established for quality control purposes and to encourage other disciplines to conduct inventories at common sampling locations. Note that:

1. The sampling point is selected with replacement. A grid point can therefore be selected more than once. Sampling points should be selected randomly with known probability. In the unlikely event that a point in the polygon is selected more than once, the information gathered at this point would be weighted accordingly.
2. If a sample polygon was selected several times, the number of grid points selected in that polygon should equal the number of times it was selected during the polygon selection process.

4.2 Process

Each selected sample polygon can be overlaid with the 100-meter grid system, which is permanent, and tied to the 20-kilometre grid of the Canadian National Forest Inventory. Copies of this grid are available from MSRM Terrestrial Information Branch in IGDS and ARC format, NAD83.

Once the grid is established on a sample polygon, there are several ways to pick the sampling point(s). Two possible methods are described:

1. Select a sampling point as follows:
   (i) Plot the sample polygon with the grid overlay on scale of 1:20,000, 1:10,000 or 1:5,000.
   (ii) Number the grid points in any sequence of your choice, beginning with the number 1. It is recommended to start from the top of the polygon, then go down the polygon in a zigzag pattern, working your way to the bottom, until all grid points within the polygon are numbered.
   (iii) Generate random numbers equivalent to the number of times a sample polygon was selected. The numbers should be between 1 and the number of points in the sample polygon.
   (iv) If a generated random number matches with a grid point number, that grid point is selected for ground sampling.
   (v) Record the sample number and the selected grid point.
   (vi) Record the UTM or the latitude and longitude of the grid point.

(vii) Transfer the grid point to a document photo to assist in the ground establishment of sample plots.

2. Number the grid columns (X) and rows (Y), and select a random X and a random Y. The grid intersection defined by the (X, Y) pair is the sampling point. The process is repeated if additional sampling points are needed or the selected (X, Y) pair falls outside the boundaries of the sample polygon.

   The selection of sampling points may be implemented manually or in an automated fashion (e.g. in a Geographic Information System environment). Automated selection will reduce the potential for making errors, and will improve the turn-around time for preparing field packages. Whatever approach is chosen, it is important to maintain the principles outlined in Section 4.1 and to document exactly what was done.

4.3 Small polygons

There will be some instances when a selected polygon is long and narrow, such that no 100 meter grid intersection falls in it. In such situations, a short interval system, tied to the 100 meter grid should be established. It is recommended to use a step-down approach. Start with a 50 meter interval, and if this does not work, move to a 25 meter interval, and so on.

5.0 Quality Assurance

5.1 Principles

Carefully following the procedures for sample selection is critical to ensuring the validity of a sample. To minimize selection errors, targeted data checks should be performed to detect as many errors as possible before ground sampling starts. Invalid samples cannot be used in compiling an inventory so it is important to ensure that samples are valid prior to field sampling.

5.2 Checking the population list

Some of the quality control checks that should be conducted on the population have been described in Section 2.7. In this section, a more complete list of attributes that should be checked is provided.

A check on the population list should concentrate on verifying the accuracy and completeness of the population, in terms of certain attributes. The checking involves generating population area proportions by the following attributes:

- Mapsheet
- BCLCS code (mainly Non-Vegetated, Vegetated Non-Treed, and the Vegetated Treed)
- Leading species
- Crown closure class
- Projected age class
- Projected height class
- Site index
- Operability
- PSYU
- FIZ
- Ownership
- Non-productive and non-forest descriptor codes

5.3 Checking sample distributions

After sample selection, it is important to check the relative frequency of samples against the area proportions of the attributes described in Section 5.2. In ideal situations, the sample distributions should match the population distributions very near exactly. This is particularly so for those attributes used to define the strata.

For other attributes, the match may not be exact. In such cases, an approximate 10% difference in the distributions should be considered tolerable. Differences greater than 10% or so for several attributes under investigation should trigger an immediate investigation of the selection process.

Plotting the sample locations on small-scale maps will also provide a visual check on the spatial distribution of the samples. The distribution of samples relative to the distribution of the forested landbase and biogeoclimatic zones is particularly relevant.

The sample should not be discarded or replaced unless errors in the data preparation or sample selection process are identified.

5.4 Evaluate data gaps

If overlaps and gaps occur in any project due to incompatible boundaries being combined, the proportion of area affected should be documented.

5.5 Check for effect of majority rule

The total area of the resultant polygons before the application of the majority rule should be equal to the total area after the polygon segments have been consolidated based on the majority rules (see Section 2.5). This check should be performed to ensure that polygon consolidation is not affected negatively by the majority rules.

6.0 Documentation

Good documentation facilitates more efficient data analysis and provides good background to the final report describing the various aspects of the inventory process. Sample selection documentation provides the vital link between the photo-interpretation process and the ground sampling.

6.1 Purpose

Thorough documentation of the sample selection process is required for all VRI projects. The purpose of the documentation is to ensure continuity in the sample selection process, permit auditing of the process, and serve as a reference during data analysis.

The person or people involved in completing all tasks listed below should be identified in the documentation as well as when the task was completed.

6.2 Data versions

As part of the data preparation process, the following tasks should be completed and documented:

- Confirm dates of last depletion updates.
- Establish dates of last re-inventory.
- Determine existence of ad hoc photo-interpretations.
- Maintain a list of the map/attribute file versions used for sample selection.

6.3 Features of selection process

Details of the sample selection process will vary from project to project. The selection process for each project should be described to document unique features. All issues that arise, or errors that are encountered, should be recorded and described briefly.

To achieve the documentation objectives, a working file and a final report file should be opened for each project. The working file will track the steps of the selection and contain records of anomalies encountered. The report file will contain a summary report describing the general selection process (see Section 7). This is the file that will contain the sign off of the various components of the selection process.

6.4 Stratification criteria

The stratification specifications in the sampling plans will vary depending on stakeholder needs in different inventory units. The person or people charged with sample selection should summarize and file the specifications in a one-page description. The description should contain the following information:

- Project-specific stratification criteria.
- Identify the merged strata.
- Any issues that will assist in data analysis.

6.5 Sample allocation

The characteristics of the sample allocation process dictate how the sample data should be handled in future analyses. The method used to allocate samples should be stated explicitly in the documentation.

The documentation should include the following:

- Method of sample allocation to strata.
- Identify disproportional allocation projects.
- Specify method of allocation.

6.6 Sample size

For sample selection, it is important to:

- Identify the source of sample size information.
- Provide reference to the document containing the sample size computation if applicable.

6.7 Sample and strata weights

Sample weights are computed as the total area of a stratum divided by the number of samples in the stratum.

Strata weights show the relative strata areas in relation to total land area covered by the entire population (i.e., total stratum area, divided by total area covered by target population).

Weights may be calculated based on the number of samples intended to be completed, but as the number of samples completed is a function of many other variables, it is most important to document strata areas and intended number of samples. Analysts can calculate the weights more accurately following completion of the ground sampling project.

6.8 Documentation of errors

All data errors encountered during the data preparation and sample selection processes must be documented and reported. Efforts made to correct data errors or conscience decisions to ignore data errors should also be documented.

6.9 Use of sample lists

A document explaining how the sample lists are used should be prepared and provided to the VRI project manager. This should include specific instructions on the samples to be completed, the process for dropping plots and documentation of dropped plots, the use of appropriate maps for the project, and contact information should the project manager have questions or concerns regarding the appropriate use of the sample list.

6.10 Archiving of data

Digital copies of the following items should be archived, preferably on a stable, portable and easily accessible medium, such as CD:

- Original data used for sample selection dataset assembly (all map and attribute data)
- Modified data used for sample selection (the assembled sample selection database)

- QA reports
- Sample selection software
- Sample lists
- Detailed instruction on appropriate use of the sample list (including copies of emails provided to project managers)
- Extra samples (with detailed instruction on appropriate use of the sample list)
- Maps showing plot distribution over inventory unit and plot locations within polygons
- Notes recorded during the sample selection procedure
- Copy of the final report referred to below

Note that the Ministry uses the ARCS filing name 13998-30/Project Name for all sample selection related documentation.

7.0 Reporting

7.1 Purpose

The purpose of the sample selection report is to summarize, in a few pages, the sequence of events and the specific characteristics of the selection process for each project. The report shou'd be concise, but informative. The documentation generated, as discussed in Section 6, should be used to generate a summary report.

7.2 Report contents

The sample selection report should contain the following information:

- Data assembly process
- Unit specific stratification
- Polygon and grid point selection process
- Results of quality assurance process
- Sample lists
- Selection issues
- Roles and responsibilities
- How sample lists should be used

8.0 Approval Protocols

8.1 Overview

There are three critical stages in the sample selection process when decisions must be made prior to moving on to the next stage in the VRI process. The critical stages include:

1. the validation of the sample list,
2. the validation of the selected sample, and
3. the approval of the sample selection report.

At each stage, a number of individuals should be identified who will provide approval or a sign off of preceding work.

This approval process is intended to control the quality of the selection process. The fact that an approval process is required emphasizes the importance of the sample selection exercise.

8.2 Validation of population

One or two other individuals capable of cross-checking the population list to ensure that it is reasonably complete should support the person responsible for sample selection. It is unlikely that the support people will invent new ways of evaluating the list, but their role is to ensure that all the necessary precautions are taken to construct a reasonably error-free list.

8.3 Validation of selected samples

As a standard, the sample list should always be compared against the population distribution. It is important to note that the distribution percentages are usually based on area proportion for the population. But for the sample, the percentages are based on number of samples (frequency).

A support person should also compare the population and sample distributions. This is an important check in that it is the last opportunity for the sample person and the support person to make any corrections. Once the sample selection person and support person sign off on the distributions, there is a commitment to establish the selected samples on the ground.

8.4 Approval of final report

The person responsible for completing the sample selection should prepare the sample selection report. The support person reviewing the sample selection should also review the report and sign it off.

Appendix 1 - Glossary

This section provides a glossary of the terms used in the manual.

Accuracy	The success in estimating the true value of an attribute.
Attribute Data	Numeric or descriptive data for each polygon.
BEC - Biogeoclimatic Zones	A geographic area having similar patterns of energy flow, vegetation and soils as a result of a broadly homogenous macroclimate.
Crown Closure	The percentage of ground area covered by the vertically projected crowns of shrubs or trees.
Data Source	The primary source of information used to determine the attribute or attributes being described.
GIS – Geographic Information System	Computer system designed to allow users to collect, manage and analyze large volumes of spatially referenced information and associated attribute data.
Graphic Data	Digital information showing polygons such as forest cover maps and recreation maps.
Ground Sampling	Taking observations (samples) on the ground.
Inventory Unit	The population of interest, usually a large-scale management unit such as a TSA or TFL.
Majority Rules	A set of rules designed to reduce the presence of numerous small resultant polygons being associated with a single polygon attribute by assigning to that polygon the overlay attribute that makes up the greatest percentage of the area.
Operability	That portion of the production forest that, under current market conditions, can be harvested at a profit.
Operability Lines	A line drawn on a map to differentiate between areas that are operable and those that are not, given status quo harvesting and reforestation technology. The operability line is used to determine the operable land base in long-run, sustained yield calculations.
Overlay	A level of information in a GIS that provides polygon information for one attribute, such as ownership, forest cover, or operability. Overlays containing different attribute information are merged together to form resultant polygons.
Polygon	Delineations or enclosures that represent discrete areas on a map or in a GIS, bounded by a line.
Population	The aggregate of all non-overlapping attributes for the inventory unit for which you want information.
Population Maps	A map of all resultant polygons in the population

PPS	Sampling design where the probability of selection is proportional to size (polygon area), so that the larger polygons have a proportionately greater chance of being included in the sample relative to smaller polygons.
PPSWR	Sampling design where the probability of selection is proportional to size (polygon area) with replacement of samples back into the sampling frame, allowing for the possibility of one sample being selected more than once.
Probability	The likelihood of a particular uncertain event occurring.
PSYU – Public Sustained Yield Unit	A portion of a TSA. An area of Crown land, usually a natural topographic unit determined by drainage areas, managed for sustained yield.
Resultant Polygon	Polygons that result after the merger of different overlays of information.
Sample	Part of the population, used to gain knowledge about the population.
Sample Size	The number of samples selected in an inventory unit.
Sample Weight	Total area of a stratum divided by the number of samples in the stratum. Shows the relative strata areas in relation to total land area covered by the entire population.
Sampling Frame	The complete list of samples selected.
Sampling Method	The method of selecting samples from the sampling frame.
Sampling Objectives	A statement indicating the geographic area and design specifications to be used to develop the sampling plan, such as the key attributes and the desired precision of results.
Sampling Point	A point on the ground inside the sample polygon.
Selection List	A sorted list of all population elements eligible for selection as samples.
Site Index	An expression of the forest site quality of a stand, at a specified age, based either on the site height, or on the top height, whichever is a more objective measure.
Strata	Sub-unit of the population which has been segmented by an attribute.
Stratification	The process of dividing the inventory unit into sub-populations called strata.
Stratification Attributes	The specific attributes that the population is to be segmented by.

Appendix 2 - Sample List Attributes

Data Management	VRI Project #	most important where projects overlap or are combined to provide overall answers - does not change - assigned by MSRM, TIB
	Sample #	should never change
	Map #	should never change
	Poly #	polygon number identified in the sample list - linked to map version
	Map Version	the graphics version used in the production of the sample list - this is important in determining the population sampled, and that the correct inventory and ground data are used in analysis
	New Map Version	if new maps are to be used for analysis, or are used for identification of the sample location within the poly
	Polygon Area	required
	New Poly #	poly # may change if map updated or reinventoried following the preparation of the sample list
	Sample IPC UTM Coordinate	either from map or GPS - useful especially when overlaying sample locations on new forest covers or other overlays, should factor in any GPS offsets
	UTM Coordinate Source	map, GPS
	Sample Type	same code as card
	Dropped - Inaccessible	samples in population but cannot be accessed, why dropped
	Inaccessible Replacement map/poly/version/coord	needed to ensure that the correct inventory data is used in the analysis
	Dropped - outside of population	plot location does not meet project population criteria - ie: plot in unmapped private/park land, unmapped logged polygon, etc., why dropped
	Plot collected to make up for dropped plot where outside of population	if additional sample collected in place of dropped sample - not a replacement as in Inaccessible, just a sample correctly identified as being 'in'
	Notes to aid in analysis	field notes that describe stand characteristics that would help analysts understand why the sample data does not match the map label, why missing ages, heights, etc.
	Other Stratification Criteria	any notes that may be useful in future post-stratification of the sample - plot fell in woodlot, park, etc.

Forest Cover Attributes	all	Whether using attribute data from older inventories or VRI phase 1, it is useful to carry all forest cover attribute information forward. In the case of .fip files, culturally defined attributes and history information is also often useful

Other Attributes	TSA Number	required for some taper equations
	TSB Number	
	TFL Number	if applicable
	Update Date	may be used in analysis
	Map FIZ	required for compilation
	Map BEC	required for compilation; require zone, subzone, variant; phase required in some cases
	Map PSYU	required for compilation
	Map PSYU Block	required for compilation
	Inventory Region	optional
	Inventory Compartment	optional
	Ownership	required if used sampling frame defined by ownership values
	Forest Region	optional
	Forest District	optional
	Operability	required if used sampling frame defined by operability values
	THLB	required if used sampling frame defined by the timber harvest land base
	Management Unit Type	required if more than one management unit defined within the sampling frame
	Management Unit #	required if more than one management unit defined within the sampling frame
	Management Unit Name	required if more than one management unit defined within the sampling frame
	Supply Block	optional
	TS Analysis Date	optional
	TS Analysis Unit	optional
	Sample Establishment Date	optional - example of field useful to project manager
	Field Access Notes	optional - example of field useful to project manager

Both the **Data Management** fields and **Forest Cover Attributes** fields should be included and populated in all cases. Those fields identified as being 'required for compilation' in the Other Attributes fields must be populated prior to any analysis, and are often most conveniently populated in the sample selection stage. Fields identified as 'optional' should be populated at the sample selection stage if these attributes will be used to determine either populations or will be used in future analyses.